

## **Construction of recombinant sox2-encoding plasmids that regulate pluripotency of breast cancer stem cells from indonesian patient**

**Hanifah Rahmi<sup>1</sup>, Tiodinar Theresia<sup>2</sup>, Amarila Malik<sup>3</sup>, Septelia Inawati Wanandi\*<sup>4</sup>**

<sup>1</sup> *Faculty of Pharmacy and Science, Universitas Muhammadiyah Prof. DR. HAMKA  
Islamic Center Delima Raya II/IV Street, Perumnas Klender, East Jakarta, Indonesia*

<sup>2,3</sup> *Faculty of Pharmacy, Universitas Indonesia*

*Jl. Prof. Dr. Mahar Mardjono Street, UI Campus Depok, Indonesia*

<sup>4</sup> *Department of Biochemistry & Molecular Biology, Faculty of Medicine, Universitas Indonesia*

*6<sup>th</sup> Salemba Raya Street, Jakarta, Indonesia*

*Submitted: 12-11-2018*

*Reviewed: 30-11-2018*

*Accepted: 03-05-2019*

### **ABSTRACT**

A therapy development with recombinant protein is a new and potential innovation to destroy cancer stem cells (CSCs). Various ways of killing CSCs include the provision of polyvalent anti-protein antibodies that code pluripotency. Therefore, it takes a mixture of Oct-4, c-Myc, Sox2, and Klf4 protein antigens that can stimulate the formation of polyvalent antibodies. This study aimed to construct Sox2 recombinant by identifying the target genes by reverse transcriptase PCR and then arranging their designs to be inserted into the cloning vector pET101/D-TOPO®. The target gene was developed by finding the complete sequences of Sox2 nucleotides on the NCBI GenBank. The growth on LB-ampicillin agar plates was amplified by PCR to obtain colonies with pET101/D-TOPO® vectors and inserts of the pluripotent gene of CSCs, then the PCR results were observed through electrophoresis. A total of fifteen colonies have DNA bands with a base pair of about 300 bp in length. The recombinant clones produced Sox2 genes with a base length of 330 bp.

**Keywords:** construction, pluripotency, recombinant, Sox2, Cancer Stem Cell (CSC)

---

#### **\*Corresponding author:**

Septelia Inawati Wanandi

Department of Biochemistry & Molecular Biology, Faculty of Medicine, Universitas Indonesia

Jl. Salemba Raya Street No. 6, Jakarta, Indonesia

Email: [septelia@gmail.com](mailto:septelia@gmail.com)

## INTRODUCTION

A stem cell is a cell with unique self-reproduction and renewability. Another uniqueness is pluripotency, that is, the ability to differentiate into various types of cells based on its environment (Lynch *et al.*, 2006). Cancer stem cells (CSCs) possess pluripotent characteristics expressed by major markers such as Oct-4, Sox2, c-Myc, and Klf4. These properties are associated with normal stem cells in a way that CSCs can perform self-renewal and differentiation into multiple types of cell and have a high survival rate (Shi *et al.*, 2012).

Septelia *et al.* (2011) have isolated and cultured breast CSCs from patients to compare their pluripotency levels and their association with the expression of Oct4 marker in each cell fraction. Also, due to the identification of several cell surface markers, namely CD44+, CD24-, ESA+, and CD133+, in different breast CSC subpopulations, these CSCs are classified a heterogeneous cell population (Aulia, 2012).

Molecularly, the pluripotency of stem cells is characterized by the expression of transcription factors and proteins on their surface, called cell surface markers. Based on the research conducted by Takashi K. and Yamanaka S. (2006), Sox2 is a transcription factor that promotes stem cells differentiation besides Oct4, Klf4, and c-Myc.

Sox2 has a vital role in cells determination, differentiation, and proliferation. An increase in Sox2 gene expression correlates with more severe breast cancer. Overexpression of Sox2 in MCF-7, a breast cancer line, causes cell proliferation and tumorigenesis in the G1/S phase transition of the cell cycle (Al-Dhfyar, 2013).

Sox proteins play an essential part during organogenesis and in the embryonic development of some tissues. In oncogenesis, stemness and pathways within the embryo can be reactivated. Members of Sox family have been found to regulate various tumors (Dong *et al.*, 2004). Sox2 has been detected as an immunogenic antigen in a significant percentage of lung cancer patients (Gure *et al.*, 2000) and meningioma patients (Comtesse *et al.*, 2005). In the pancreas, Sox2 expression is involved in the invasion and metastasis of pancreatic intraepithelial neoplasia (Sanada *et al.*, 2006). It has been recently reported as a marker in gastric cancer (Li *et al.*, 2006) and prostate cancer (Sattler *et al.*, 2000) and identified as a lineage-survival oncogene in squamous cell carcinoma of the cell (Hussenet *et al.*, 2010). Another investigation has also found it in the early stage of a variety of postmenopausal breast carcinomas and metastatic lymph nodes (Lengerke *et al.*, 2011).

The particular characteristics of Sox2 can be used to estimate the presence of CSCs in a group of cancer cells. In other words, Sox2 is a marker that pinpoints the location of CSCs. The coding of this pluripotency factor can be traced by generating commercially available antibodies. However, because these antibodies are manufactured to locate healthy stem cells, the production of antibodies for cancer diagnosis in Indonesia requires the cloning process of genes that code the pluripotency of cancer stem cells in Indonesian patients.

## MATERIALS AND METHOD

### Materials

The main research materials were Sox2-encoding RNAs isolated from breast cancer stem cells, which were obtained from the Cell Culture Laboratory for Cancer Stem Cells in the Department of Biochemistry FKUI. The other materials were the reaction mix kit SuperScript III First-Strand Synthesis SuperMix (Invitrogen, USA), PureLink PCR Purification Kit (Invitrogen, USA), High-Speed Plasmid Mini Kit protocol (GeneAid, Taiwan), DreamTaq PCR Master Mix (Thermo Scientific, USA), KAPA HiFi HotStart ReadyMix kit (GeneAid, Taiwan), *Escherichia coli* DH5 $\alpha$  (BPPT, Indonesia), vector pET101/D-TOPO (Invitrogen, USA), agarose (GE Healthcare, Sweden), Tris base (USB, USA), ethidium bromide, LB Broth (Sigma, Germany), LB Agar (Sigma, Germany), Bacto Agar (Becton, Dickinson & Company, France), CaCl<sub>2</sub>, glycerol (Mallinckrodt, Mexico), nuclease-free aquadest, Aqua Bidest (Brataco, Indonesia), and Aqua Dest (Brataco, Indonesia).

### Primer Design

The primer was designed using Serial Cloner and PerlPrimer programs. The complete sequences of mRNA and amino acid of the pluripotency marker Sox2 were obtained from the NCBI GenBank. Then, the DNA binding site was obtained from the UniProt website. Open reading frames (ORFs) of mRNA and active domain position of the sequence were determined with Serial Cloner. Using PerlPrimer, the research was able to produce a variety of primer pairs, i.e., forward and reverse primer, and then select the compatible primer. Afterward, an alignment was conducted between selected primers and the mRNA sequences ([Marshall, 2004](#)).

### Reverse transcriptase-PCR

RT-PCR was executed using the SuperScript® III One-Step RT-PCR System with Platinum® Taq DNA Polymerase. The RT-PCR component consisted of 25 µl of 2x reaction mix, 1 µl of RNA (concentration= 300 ng), 1 µl of forward primer (concentration= 10 µM), 1 µl of reverse primer (concentration= 10 µM), 2 µl of SuperScript III Reverse Transcriptase, and 50 µl of nuclease-free water. The RT-PCR conditions were as follows: cDNA synthesis was carried out at 55°C for 30 minutes; pre-denaturation at 94°C for 2 minutes; 40 cycles, each consisting of denaturation at 94°C for 15 minutes, annealing at 60°C (Sox2), and extension at 68°C for a minute; post extension at 68°C for 5 minutes ([Wacker and Godard, 2005](#)).

### Polymerase chain reaction

The RT-PCR purification product was amplified using KAPA HiFi HotStart ReadyMix kit. Mixtures of 9 µL of ddH<sub>2</sub>O, 0.75 µL of forward primer, 0.75 µL of reverse primer, 2µL of DNA template, and 12.5 µL of KAPA HiFi ready mix were placed in a 0.5 mL microtube. The PCR conditions were as follows: pre-denaturation at 95°C for 5 minutes; 35 cycles, each consisting of denaturation at 98°C for 30 seconds, annealing at 60°C (Sox2 and Klf4) for 15 seconds, extension at 72°C for 1 minutes; post extension at 72°C for 5 minutes ([Mathieu et al., 2014](#)).

### pET101/D-TOPO vector cloning

The cloning reaction was completed by putting the following components into a microtube: 4µl of fresh PCR product; 1 µl of salt solution; 1µl of pET TOPO® vector. Then, the microtube was incubated on ice, and the mixed components transformed into DH5α competent *E. coli* ([Mathieu et al., 2014](#)).

### Production of competent cells

One colony of *E. coli* was taken from the stock, incubated for 16-20 hours at 37°C, and transferred into a test tube containing 10 ml of LB broth, then incubated overnight at 37°C with stirring. Subsequently, 0.4 ml of culture was injected into 10 ml of new and sterile LB liquid medium and incubated for 3 hours at 37°C with stirring. Aseptically, 1.5 mL of culture was inserted into a sterile microtube and centrifuged at 5000 rpm for a minute in a refrigerated microcentrifuge. The cells pellet was washed with 1 mL of cold 0.1M CaCl<sub>2</sub> and set aside on ice for 15 minutes. The centrifugation was conducted at 5000 rpm for one minute, and the cells pellet was suspended in 200 µL of cold CaCl<sub>2</sub> solution, then set aside on ice for 10 minutes. Afterward, the competent cells were ready to use ([Mathieu et al., 2014](#)).

### Transformation of target gene

One hundred µL of sterile 0.1M CaCl<sub>2</sub> was inserted into a sterile microtube and added with a sample of the ligation product and 200 µL of DH5α competent *E. coli*. For the control, the sample was replaced with ddH<sub>2</sub>O. Then, it was incubated on ice for 90 minutes and stirred with the index finger every 10 minutes. A heat shock was delivered at 42°C for 90 seconds without shaking. The tube was transferred to the ice for 2 minutes, added with 250µl of SOC medium at room temperature, tightly

sealed, and shaken horizontally at 200rpm, 37°C, for 90 minutes. The transformed suspension was spread on the LB agar medium containing ampicillin and incubated overnight at 37°C. The grown colonies were selected and replicated by streaking and picking the colonies then placing them into a new LB-ampicillin agar. These selected colonies were incubated overnight at 37°C. The replica was used for further analysis through plasmid isolation (Mathieu *et al.*, 2014).

### Analysis of transformation

The plasmid was isolated using the High-Speed Plasmid Mini Kit Protocol. The Sox2-DNA association was amplified with the KAPA HiFi HotStart ReadyMix kit. Mixtures of 9 µL of ddH<sub>2</sub>O, 0.75 µL of forward primer, 0.75 µL of reverse primer, and 2 µL of DNA template were put into a 0.5 mL microtube. Then, 12.5 µL of KAPA HiFi ready mix was added into the mixture and centrifuged temporarily. The PCR conditions were as follows: pre-denaturation at 95°C for 5 minutes; 35 cycles, each consisting of denaturation at 98°C for 30 seconds, annealing at 60°C for 15 seconds, and extension at 72°C for one minute; post extension at 72°C for 5 minutes (Mathieu *et al.*, 2014).

### Sequencing

The sequence mapping was prepared to identify the sequence of the base pairs of Sox2. A verification was then performed to evaluate the absence or presence of a mutation. The sequencing reaction used the DNA Sequencing Service provided by the 1st BASE. The DNA sequencing results were analyzed in BioEdit 7.0.5.3 software (Tippmann, 2004).

## RESULT AND DISCUSSION

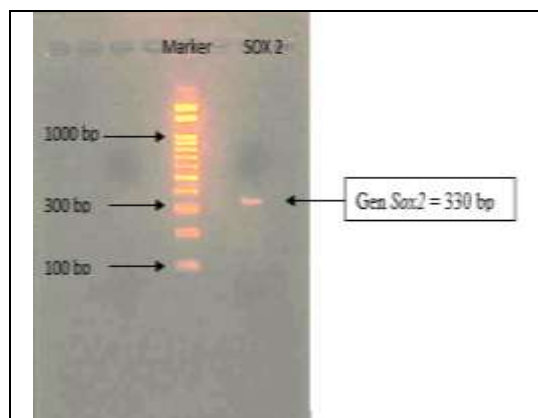
### Design of Sox2 primer

The Sox2 primer (NM\_003106.3) was designed from the nucleotide base in the 558-737th order with the sequence of the forward primer for Sox2 as follows: 5'-CACCATGAACATGATGGAGACGGAGCT-3'. Meanwhile, the sequence of the reverse primer was 5'-GGGCCGGTATTTATAATCCGG-3'. The length of the Sox2 product generated from this primer pair was 330 bp.

The sequence encoding the DNA-binding domain was input to the PerlPrimer program by adding triplicate bases at the beginning and the end of the strand of the sequence. The forward primer was added with seven bases of CACC ATG to adjust the cloning strategy. In this case, the addition of CACC was to allow the overhang of the Sox2 genes that would be inserted at the 5' end of the cloning vector (GTGG), while ATG was added to initiate the formation of the start codon during the transcription process so that the protein could be expressed. The addition of CACC so that there is an overhang between the Sox2 gene which will be inserted into the 5' end of the cloned vector (GTGG) while ATG is added so that the start codon is formed so the protein can be expressed.

### Reverse transcriptase-polymerase chain reaction

The Sox2 RT-PCR products were collected with 2% gel electrophoresis, as visualized in Figure 1. The RT-PCR reaction was performed using the SuperScript® III One-Step RT-PCR System with Platinum® Taq DNA Polymerase. The annealing temperature was adjusted to the melting temperature of the primer. The RT-PCR product of the Sox2 gene had a single band located by the benchmark length of the Sox2 product, i.e., 330 bp.



**Figure 1. The RT-PCR analysis of Sox2**

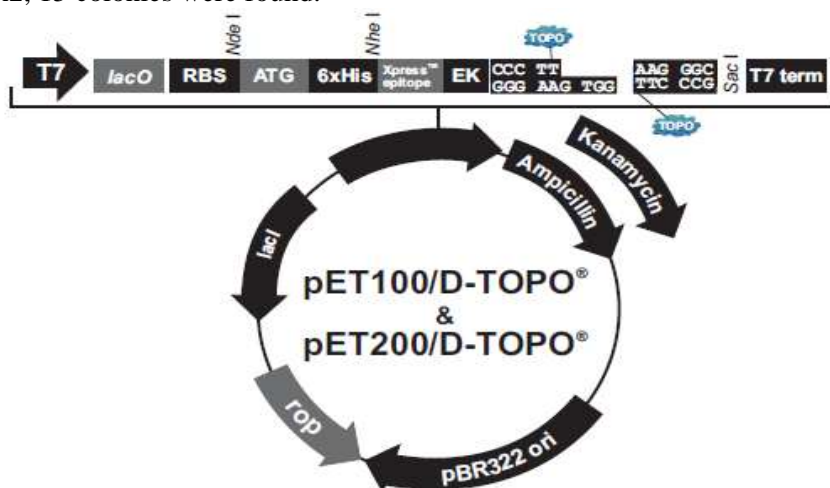
### Sequencing of PCR product

The DNA of Sox2 was amplified using KAPA HiFi HotStart ReadyMix Kit. The PCR products were analyzed by sequencing to ensure that the formed bands were the genes of Sox2. Sequencing was conducted using the DNA Sequencing Service provided by 1st BASE using two reactions, i.e., forward and reverse. Based on the sequencing results of the Sox2 gene, the nucleotide bases were aligned using the Basic Local Alignment Search Tool (BLAST) on the NCBI website.

The analysis results of the nucleotide-nucleotide BLAST (BLASTn) of the forward sequence of Sox2 showed 94% identity with the Sox2 sequence with the access code NM\_003106.3, while the reverse one demonstrated 93% identity with the Sox2 sequence with the access code NM\_003106.3.

### Transformation of Target Gene

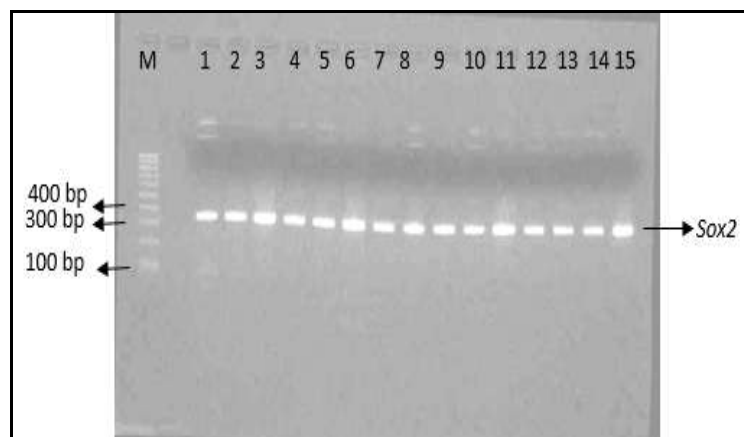
The target gene obtained in this research was the result of PCR with a blunt end. The growth of several colonies was detected in the agar medium but unidentified in the control medium. In the agar medium of Sox2, 15 colonies were found.



**Figure 2. pET100/D-TOPO® vector**

The colonies that grew on the LB-ampicillin medium had been inserted with the pET101 vector (Figure 2). This vector has a site that is resistant to ampicillin; hence, the colony containing this vector was able to survive on the LB-ampicillin medium. However, these colonies require more analysis

concerning the successful transformations with PCR colonies. The results of the PCR plasmids of Sox2 suggested that all isolated plasmids comprised Sox2 (330 bp) genes (Figure 3).



**Figure 3.** The PCR results of the colony from the insertion of the Sox2 gene (the annealing temperature= 60°C)

### Sequencing of PCR colonies

The PCR results of the colonies were analyzed using the same procedure as the sequencing of the PCR Product, that is sequencing and the use of 1<sup>st</sup> BASE and BLAST software to define the nitrogenous bases of the Sox2 gene. The BLASTn analysis results depicted 92% identity with the Sox2 sequence with the access code NM\_003106.3 (Figure 4), meaning that the sequence of the colonies matches or shares 92% similarity with human Sox2 sequence in the NCBI data bank (Barton, 2007).

Download v GenBank Graphics					
Homo sapiens SRY-box 2 (SOX2), mRNA					
Sequence ID: <a href="#">NM_003106.3</a> Length: 2520 Number of Matches: 1					
Range 1: 444 to 772 GenBank Graphics					
▼ Next Match ▲ Previous Match					
Score	Expect	Identities	Gaps	Strand	
475 bits(257)	3e-131	308/335(92%)	6/335(1%)	Plus/Plus	
Query 379	AACATGATGGAGACGGAGCTGAAGCCGYCKKRCYCGCWGCAAGCTTCGGGGGGCGGCGGC				438
Sbjct 444	AACATGATGGAGACGGAGCTGAAGCCGCCGGGCCCGCAGCAAACTTCGGGGGGCGGCGGC				503
Query 439	GGAGGAGGTAACGCCACGGCGGCAGCGACCGGCGGCAACCAGAAGAAGAGCCCGGACCGC				498
Sbjct 504	-----GGCAACTCCACCGCGGCGGCGCGGCGGCAACCAGAAAAACAGCCCGGACCGC				557
Query 499	GTCAAGAGGCCCATGAACGCCTTCATGGTGTGGTCCCGGGGCGAGCGGCGTAAGATGGCC				558
Sbjct 558	GTCAAGCGGCCCATGAATGCCTTCATGGTGTGGTCCCGGGGCGAGCGGCGCAAGATGGCC				617
Query 559	CAGGAGAACCCCAAGATGCACAACCTCGGAGATCAGCAAGCGCCTCGGCGCCGAGTGGAAA				618
Sbjct 618	CAGGAGAACCCCAAGATGCACAACCTCGGAGATCAGCAAGCGCCTGGGCGCCGAGTGGAAA				677
Query 619	CTTTTGTGCGGAGACCGAGAAGCGGCCGTTTCATCGACGAGGCCAAGCGGCTGCGAGCTCTG				678
Sbjct 678	CTTTTGTGCGGAGACCGAGAAGCGGCCGTTTCATCGACGAGGCTAAGCGGCTGCGAGCGCTG				737
Query 679	CACATGAAGGAGCACCCGGATTATAAATACCGGCC		713		
Sbjct 738	CACATGAAGGAGCACCCGGATTATAAATACCGGCC		772		

**Figure 4.** The alignment analysis of the sequencing



## CONCLUSION

The designed forward and reverse primers have successfully amplified the active domain of Sox2, which controls the gene that codes pluripotency. The Sox2 genes have been successfully cloned to pET101/D-TOPO vectors and transformed into *Escherichia coli* DH5a host. The clone can be further developed to generate antibodies that can detect a cancer cell in Indonesian cancer patients.

## ACKNOWLEDGEMENT

This work was supported by the Research Grant (RISBIN IPTEKDOK) 2013 managed by the Ministry of Health of Indonesia which had announced at number LB.02.01/I/1/891/2013.

## REFERENCES

- Al-Dhfyhan A., 2013. Embryonic signature in breast cancers; pluripotency roots of cancer stem cells, *Saudi Pharmaceutical Journal*, 21(2): 229–232.
- Aulia G., 2012. Modulasi stress oksidatif pada ketahanan hidup sel punca kanker payudara (CD24-/CD44+) melalui Induksi Rotenon: tinjauan pada ekspresi MnSOD, *Master Thesis*, Jakarta: Fakultas Kedokteran Universitas Indonesia.
- Barton G. J., 2007. Sequence alignment for molecular replacement, *Acta Crystallographica Section D Biological Crystallography*, D64: 25-32.
- Comtesse N., Zippel A., Walle S., Monz D., Backes C., Fischer U., Mayer J., Ludwig N., Hildebrandt A., Keller A., Steudel W. I., Lenhof H. P., and Meese E., 2005. Complex humoral immune response against a benign tumor: frequent antibody response against specific antigens as diagnostic targets. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27): 9601–9606.
- Dong C., Wilhelm D., and Koopman P., 2004. Sox genes and cancer, *Cytogenetic and Genome Research*, 105(2-4): 442–447.
- Gure A. O., Stockert E., Scanlan M. J., Keresztes R. S., Jager D., Altorki N. K., Old L. J., and Chen Y. T., 2000. Serological identification of embryonic neural proteins as highly immunogenic tumor antigens in small cell lung cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 97(8): 4198–4203.
- Hussenet T., Dali S., Exinger J., Monga B., Jost B., Dembele D., Martinet N., Thibault C., Huelsken J., Brambilla E. and Manoir S., 2010. SOX2 Is an Oncogene Activated by Recurrent 3q26.3 Amplifications in Human Lung Squamous Cell Carcinomas. *Plos One*, 5(1): e8960.
- Lengerke C., Fehm T., Kurth R., Neubauer H., Scheble V., Müller F., Schneider F., Petersen K., Wallwiener D., Kanz L., Fend F., Perner S., Bareiss P. M., and Staebler A., 2011. Expression of the embryonic stem cell marker SOX2 in early-stage breast carcinoma, *BMC Cancer*, 11(42): 1-10.
- Li X. L., Eishi Y. B., Bai Y. Q., Sakai H., Akiyama Y., Tani M., Takizawa T., Koike M. and Yuasa Y., 2004. Expression of the SRY-Related HMG Box Protein SOX2 in human gastric carcinoma, *International Journal of Oncology*, 24(2): 257–263.
- Lynch M. D., Cariati M., and Purushotham A. D., 2006. Breast Cancer: Stem Cells and Prospects for Therapy, *Breast Cancer Research*, 8(3): 211.
- Marshall O. J., 2004. PerlPrimer: Cross-Platform, Graphical Primer Design for Standard, Bisulphite and Real-Time PCR, *Bioinformatics*, 20(15): 2471-2472.
- Mathieu J., Alvarez E., and Alvarez P., 2014. Recombination-Assisted Megaprimer (RAM) Cloning, *MethodsX*, 1: 23-29.
- Sanada Y., Yoshida K., Ohara M., Oeda M., Konishi K., and Tsutani, 2006. Histopathologic evaluation of stepwise progression of pancreatic carcinoma with immunohistochemical analysis of gastric epithelial transcription factor SOX2 - comparison of expression patterns between invasive components and cancerous or nonneoplastic intraductal components. *Pancreas*. 32(2): 164–170.

- Sattler H. P., Lensch R., Rohde V., Zimmer E., Meese E., Bonkhoff H., Retz M., Zwergel T., Bex A., Stoeckle M., and Wullich B, 2000. Novel amplification unit at chromosome 3q25-q27 in human prostate cancer, *Prostate*, 45(3):207–215.
- Septelia I.W., Reni P., Naroeni A., Sonar S. P., Endang W. B., Amarila M., and Budiman B, 2011. Characterization of manganese superoxide dismutase expression (mnsod) in human breast cancer stem cells and its correlation with cell pluripotency, *Medicine & Health*, 6(1): 208.
- Shi Z., Bai R., Fu Z., Zhu Y., Wang R., and Zheng S, 2012. Induced pluripotent stem cell-related genes influence biological behavior and 5-fluorouracil sensitivity of colorectal cancer cells, *J Zhejiang Univ Sci B*, 13(1): 11–19.
- Takahashi K. and Yamanaka S, 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors, *Cell*, 126(4): 663-676.
- Tippmann H., 2004, Software review: analysis for free: comparing programs for sequence analysis, *Briefings in Bioinformatics*, 5(1): 82-87.
- Wacker M. and Godard M., 2005. Analysis of One-Step and Two-Step Real-Time RT-PCR Using SuperScript III, *J Biomol Tech*, 16(3): 266–271.